

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/587,085  
Source: IFWP  
Date Processed by STIC: 08/03/2006

# ***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 08/03/2006

PATENT APPLICATION: US/10/587,085

TIME: 09:14:07

Input Set : A:\5404-161 Sequence Listing.txt

Output Set: N:\CRF4\08032006\J587085.raw

3 <110> APPLICANT: KANEKA CORPORATION  
 5 <120> TITLE OF INVENTION: PORYPEPTIDE HAVING AMIDASE ACTIVITY AND GENE THEREOF  
 7 <130> FILE REFERENCE: 5404-161(B040008W001)  
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/587,085  
 C--> 9 <141> CURRENT FILING DATE: 2006-07-21  
 9 <150> PRIOR APPLICATION NUMBER: PCT/JP2005/000951  
 10 <151> PRIOR FILING DATE: 2005-01-26  
 12 <150> PRIOR APPLICATION NUMBER: JP2004-028041  
 13 <151> PRIOR FILING DATE: 2004-02-04  
 15 <160> NUMBER OF SEQ ID NOS: 11  
 17 <170> SOFTWARE: PatentIn version 3.1  
 19 <210> SEQ ID NO: 1  
 20 <211> LENGTH: 388  
 21 <212> TYPE: PRT  
 22 <213> ORGANISM: Arthrobacter sp.  
 24 <400> SEQUENCE: 1  
 25 Met Ser Arg Leu Leu Arg Glu His Gly Ile Val Ile Gly Arg Leu Gln  
 26 1 5 10 15  
 28 Pro Gly Ser Leu Asn Thr Ile Ala Asp Val Ala Gly Val Arg Val Gly  
 29 20 25 30  
 31 His Ser Thr Ile Met Arg Gly Ser Gly Pro Leu Ser Ile Gly His Gly  
 32 35 40 45  
 34 Pro Val Arg Thr Gly Val Thr Ala Ile Ile Pro His Glu Gly Asp Ile  
 35 50 55 60  
 37 Trp Glu Glu Pro Arg Phe Ala Gly Val Phe Ser Leu Asn Gly Ser Gly  
 38 65 70 75 80  
 40 Glu Trp Ser Gly Thr Ser Phe Val Arg Glu Thr Gly Cys Leu Tyr Gly  
 41 85 90 95  
 43 Pro Ile Met Thr Thr Asn Ser His Ser Ile Gly Ser Val Arg Asn Ala  
 44 100 105 110  
 46 Val Ile Lys Arg Glu Val Ala Arg Arg Gly Ser Leu Glu Arg Leu Pro  
 47 115 120 125  
 49 Leu Val Gly Glu Thr Phe Asp Gly Leu Leu Asn Asp Ile Ser Gly Met  
 50 130 135 140  
 52 His Val Lys Asp Glu His Val Ala Glu Ala Ile Asp Ser Ala Ser Ala  
 53 145 150 155 160  
 55 Asn Val Thr Glu Gly Asn Val Gly Gly Gly Thr Gly Asn Val Cys His  
 56 165 170 175  
 58 Gly Phe Lys Gly Gly Ile Gly Ser Ala Ser Arg Val Leu Gln Leu Gly  
 59 180 185 190  
 61 Glu Glu Thr Tyr Thr Leu Gly Val Leu Val Gln Ala Asn His Gly Leu  
 62 195 200 205  
 64 Arg Asp Glu Phe Gln Val Thr Gly Val Pro Val Gly Arg Leu Ile Ser

Cp9-6)

## RAW SEQUENCE LISTING

DATE: 08/03/2006

PATENT APPLICATION: US/10/587,085

TIME: 09:14:07

Input Set : A:\5404-161 Sequence Listing.txt

Output Set: N:\CRF4\08032006\J587085.raw

```

65      210      215      220
67 Thr Asp Glu Ile Pro Leu Gly Pro Ser Gly Phe Asp Arg Arg Ser Ser
68 225      230      235      240
70 Pro His Lys Asn Ser Ile Leu Val Val Val Ala Thr Asp Ala Pro Leu
71      245      250      255
73 Leu Pro Gly Gln Leu Glu Arg Val Ala His Arg Ser Thr Leu Gly Ile
74      260      265      270
76 Ala Arg Asn Gly Ala Tyr Ala His Asn Leu Ser Gly Asp Leu Ala Leu
77      275      280      285
79 Ala Phe Ser Thr Cys Pro Gln Pro Val Ser Gly Tyr Asp Phe Gly Val
80      290      295      300
82 Asp Thr Ser Pro Gly Thr Ile Arg Ala Leu Pro Asn Ala Ala Thr Ala
83 305      310      315      320
85 Gly Leu Phe Glu Ala Ala Val Glu Ala Thr Glu Glu Ala Ile Val Ser
86      325      330      335
88 Ala Leu Val His Ala Asp Thr Cys Thr Gly Ile Asp Asp Arg Val Ala
89      340      345      350
91 Tyr Gly Leu Glu Ala Ala Arg Leu Ala Arg Ser Ile Ser Glu Tyr Arg
92      355      360      365
94 Gly Thr Gln Leu Tyr Pro Glu Lys Val Ser Asp Ser His Leu Glu Arg
95      370      375      380
97 Arg Ser Gln Pro
98 385
101 <210> SEQ ID NO: 2
102 <211> LENGTH: 1290
103 <212> TYPE: DNA
104 <213> ORGANISM: Arthrobacter sp.
106 <220> FEATURE:
107 <221> NAME/KEY: CDS
108 <222> LOCATION: (53)..(1219)
109 <223> OTHER INFORMATION: sequence coding for amidase
111 <400> SEQUENCE: 2
112 agcgcgtcgt ggactgggtg cagaaataca caggcgagcc cgaggacgaa aa atg agc      58
113                                     Met Ser
114                                     1
116 cgt ctg ctc cgt gag cac gga atc gtc atc ggt cgt ctc caa ccg ggt      106
117 Arg Leu Leu Arg Glu His Gly Ile Val Ile Gly Arg Leu Gln Pro Gly
118      5      10      15
120 tct ctg aac acc att gca gac gtc gca ggc gtt cgc gta ggc cat tca      154
121 Ser Leu Asn Thr Ile Ala Asp Val Ala Gly Val Arg Val Gly His Ser
122      20      25      30
124 aca atc atg cgc ggt tct ggg ccc ctg tcc atc ggc cat ggc cca gtt      202
125 Thr Ile Met Arg Gly Ser Gly Pro Leu Ser Ile Gly His Gly Pro Val
126 35      40      45      50
128 cgc aca ggg gta aca gcc atc atc cct cac gaa gga gac atc tgg gag      250
129 Arg Thr Gly Val Thr Ala Ile Ile Pro His Glu Gly Asp Ile Trp Glu
130      55      60      65
132 gag cca cgg ttc gcc ggc gtc ttc tcc ctg aat ggc agc ggt gaa tgg      298
133 Glu Pro Arg Phe Ala Gly Val Phe Ser Leu Asn Gly Ser Gly Glu Trp

```

## RAW SEQUENCE LISTING

DATE: 08/03/2006

PATENT APPLICATION: US/10/587,085

TIME: 09:14:07

Input Set : A:\5404-161 Sequence Listing.txt

Output Set: N:\CRF4\08032006\J587085.raw

134	70	75	80	
136	agc gga acc tcg ttc gtc agg gag acc ggg tgt ctt tat ggc cct atc	346		
137	Ser Gly Thr Ser Phe Val Arg Glu Thr Gly Cys Leu Tyr Gly Pro Ile			
138	85 90 95			
140	atg acg acg aat tcg cac agc att gga tcg gtc agg aac gcc gtc atc	394		
141	Met Thr Thr Asn Ser His Ser Ile Gly Ser Val Arg Asn Ala Val Ile			
142	100 105 110			
144	aag cgt gaa gta gcc cgg cgg gga agc ctg gag agg ctc cct ctc gtg	442		
145	Lys Arg Glu Val Ala Arg Arg Gly Ser Leu Glu Arg Leu Pro Leu Val			
146	115 120 125 130			
148	gga gaa acc ttt gat ggc cta ctc aat gac atc agc ggc atg cac gtc	490		
149	Gly Glu Thr Phe Asp Gly Leu Leu Asn Asp Ile Ser Gly Met His Val			
150	135 140 145			
152	aag gac gaa cac gtg gcc gag gcc atc gac tcc gcc tcc gca aat gtt	538		
153	Lys Asp Glu His Val Ala Glu Ala Ile Asp Ser Ala Ser Ala Asn Val			
154	150 155 160			
156	acc gaa ggc aat gtt ggc ggt ggg acc gga aat gtt tgt cac ggt ttc	586		
157	Thr Glu Gly Asn Val Gly Gly Thr Gly Asn Val Cys His Gly Phe			
158	165 170 175			
160	aaa ggc ggt att gga agt gcc tcg cgc gtg ttg caa ttg ggc gag gaa	634		
161	Lys Gly Gly Ile Gly Ser Ala Ser Arg Val Leu Gln Leu Gly Glu Glu			
162	180 185 190			
164	acc tac act ttg ggg gtt ctc gtc caa gcc aac cac ggc ctt cgt gac	682		
165	Thr Tyr Thr Leu Gly Val Leu Val Gln Ala Asn His Gly Leu Arg Asp			
166	195 200 205 210			
168	gaa ttt cag gtg acg gga gta ccc gtg gga agg ctc ata tct acg gac	730		
169	Glu Phe Gln Val Thr Gly Val Pro Val Gly Arg Leu Ile Ser Thr Asp			
170	215 220 225			
172	gag atc ccc ttg ggg cct tca ggt ttt gat cga agg tct tca cca cac	778		
173	Glu Ile Pro Leu Gly Pro Ser Gly Phe Asp Arg Arg Ser Ser Pro His			
174	230 235 240			
176	aaa aac agt att ctt gtc gtc gtg gca acc gac gcg cct cta cta ccg	826		
177	Lys Asn Ser Ile Leu Val Val Ala Thr Asp Ala Pro Leu Leu Pro			
178	245 250 255			
180	ggc caa ctg gaa cgc gtt gct cac cgt tct acc cta ggc att gcc cgt	874		
181	Gly Gln Leu Glu Arg Val Ala His Arg Ser Thr Leu Gly Ile Ala Arg			
182	260 265 270			
184	aat ggt gcc tac gcg cac aat ctc agc ggc gac ctt gca ctt gcg ttc	922		
185	Asn Gly Ala Tyr Ala His Asn Leu Ser Gly Asp Leu Ala Leu Ala Phe			
186	275 280 285 290			
188	tcc acc tgc ccg cag cct gta agc ggt tac gat ttc gga gtg gat aca	970		
189	Ser Thr Cys Pro Gln Pro Val Ser Gly Tyr Asp Phe Gly Val Asp Thr			
190	295 300 305			
192	agt cct ggg acc att cgc gcc ctg ccc aac gcc gca acg gct ggc ctc	1018		
193	Ser Pro Gly Thr Ile Arg Ala Leu Pro Asn Ala Ala Thr Ala Gly Leu			
194	310 315 320			
196	ttc gag gcg gcc gtt gag gcc act gag gaa gcg att gtt tcc gcg ctt	1066		
197	Phe Glu Ala Ala Val Glu Ala Thr Glu Glu Ala Ile Val Ser Ala Leu			
198	325 330 335			

## RAW SEQUENCE LISTING

DATE: 08/03/2006

PATENT APPLICATION: US/10/587,085

TIME: 09:14:07

Input Set : A:\5404-161 Sequence Listing.txt

Output Set: N:\CRF4\08032006\J587085.raw

```

200 gtc cac gcc gac acc tgc acc ggg atc gat gac agg gtt gcc tat ggg      1114
201 Val His Ala Asp Thr Cys Thr Gly Ile Asp Asp Arg Val Ala Tyr Gly
202      340      345      350
204 ttg gag gcg gct cga ctt gct cgt tca att tcg gaa tat cga ggc acc      1162
205 Leu Glu Ala Ala Arg Leu Ala Arg Ser Ile Ser Glu Tyr Arg Gly Thr
206 355      360      365      370
208 cag ctg tat ccg gag aaa gtg tcg gat tcc cat ctt gaa cga agg agc      1210
209 Gln Leu Tyr Pro Glu Lys Val Ser Asp Ser His Leu Glu Arg Arg Ser
210      375      380      385
212 cag ccg tga ccgccgcgca gccaaagccaa gcaccacccc gggcaaaggc      1259
213 Gln Pro
216 cgggaaacgg tccaacctaa cacgcaacga t      1290
220 <210> SEQ ID NO: 3
221 <211> LENGTH: 1167
222 <212> TYPE: DNA
223 <213> ORGANISM: Arthrobacter sp.
225 <220> FEATURE:
226 <221> NAME/KEY: CDS
227 <222> LOCATION: (1)..(1167)
228 <223> OTHER INFORMATION: sequence coding for amidase
230 <400> SEQUENCE: 3
231 atg agc cgt ctg ctc cgt gag cac gga atc gtc atc ggt cgt ctc caa      48
232 Met Ser Arg Leu Leu Arg Glu His Gly Ile Val Ile Gly Arg Leu Gln
233 1      5      10      15
235 ccg ggt tct ctg aac acc att gca gac gtc gca ggc gtt cgc gta ggc      96
236 Pro Gly Ser Leu Asn Thr Ile Ala Asp Val Ala Gly Val Arg Val Gly
237      20      25      30
239 cat tca aca atc atg cgc ggt tct ggg ccc ctg tcc atc ggc cat ggc      144
240 His Ser Thr Ile Met Arg Gly Ser Gly Pro Leu Ser Ile Gly His Gly
241      35      40      45
243 cca gtt cgc aca ggg gta aca gcc atc atc cct cac gaa gga gac atc      192
244 Pro Val Arg Thr Gly Val Thr Ala Ile Ile Pro His Glu Gly Asp Ile
245      50      55      60
247 tgg gag gag cca cgg ttc gcc ggc gtc ttc tcc ctg aat ggc agc ggt      240
248 Trp Glu Glu Pro Arg Phe Ala Gly Val Phe Ser Leu Asn Gly Ser Gly
249 65      70      75      80
251 gaa tgg agc gga acc tcg ttc gtc agg gag acc ggg tgt ctt tat ggc      288
252 Glu Trp Ser Gly Thr Ser Phe Val Arg Glu Thr Gly Cys Leu Tyr Gly
253      85      90      95
255 cct atc atg acg acg aat tcg cac agc att gga tcg gtc agg aac gcc      336
256 Pro Ile Met Thr Thr Asn Ser His Ser Ile Gly Ser Val Arg Asn Ala
257      100      105      110
259 gtc atc aag cgt gaa gta gcc cgg cgg gga agc ctg gag agg ctc cct      384
260 Val Ile Lys Arg Glu Val Ala Arg Arg Gly Ser Leu Glu Arg Leu Pro
261      115      120      125
263 ctc gtg gga gaa acc ttt gat ggc cta ctc aat gac atc agc ggc atg      432
264 Leu Val Gly Glu Thr Phe Asp Gly Leu Leu Asn Asp Ile Ser Gly Met
265      130      135      140
267 cac gtc aag gac gaa cac gtg gcc gag gcc atc gac tcc gcc tcc gca      480

```

## RAW SEQUENCE LISTING

DATE: 08/03/2006

PATENT APPLICATION: US/10/587,085

TIME: 09:14:07

Input Set : A:\5404-161 Sequence Listing.txt

Output Set: N:\CRF4\08032006\J587085.raw

```

268 His Val Lys Asp Glu His Val Ala Glu Ala Ile Asp Ser Ala Ser Ala
269 145          150          155          160
271 aat gtt acc gaa ggc aat gtt ggc ggt ggg acc gga aat gtt tgt cac      528
272 Asn Val Thr Glu Gly Asn Val Gly Gly Gly Thr Gly Asn Val Cys His
273          165          170          175
275 ggt ttc aaa ggc ggt att gga agt gcc tcg cgc gtg ttg caa ttg ggc      576
276 Gly Phe Lys Gly Gly Ile Gly Ser Ala Ser Arg Val Leu Gln Leu Gly
277          180          185          190
279 gag gaa acc tac act ttg ggg gtt ctc gtc caa gcc aac cac ggc ctt      624
280 Glu Glu Thr Tyr Thr Leu Gly Val Leu Val Gln Ala Asn His Gly Leu
281          195          200          205
283 cgt gac gaa ttt cag gtg acg gga gta ccc gtg gga agg ctc ata tct      672
284 Arg Asp Glu Phe Gln Val Thr Gly Val Pro Val Gly Arg Leu Ile Ser
285          210          215          220
287 acg gac gag atc ccc ttg ggg cct tca ggt ttt gat cga agg tct tca      720
288 Thr Asp Glu Ile Pro Leu Gly Pro Ser Gly Phe Asp Arg Arg Ser Ser
289 225          230          235          240
291 cca cac aaa aac agt att ctt gtc gtc gtg gca acc gac gcg cct cta      768
292 Pro His Lys Asn Ser Ile Leu Val Val Ala Thr Asp Ala Pro Leu
293          245          250          255
295 cta ccg ggc caa ctg gaa cgc gtt gct cac cgt tct acc cta ggc att      816
296 Leu Pro Gly Gln Leu Glu Arg Val Ala His Arg Ser Thr Leu Gly Ile
297          260          265          270
299 gcc cgt aat ggt gcc tac gcg cac aat ctc agc ggc gac ctt gca ctt      864
300 Ala Arg Asn Gly Ala Tyr Ala His Asn Leu Ser Gly Asp Leu Ala Leu
301          275          280          285
303 gcg ttc tcc acc tgc ccg cag cct gta agc ggt tac gat ttc gga gtg      912
304 Ala Phe Ser Thr Cys Pro Gln Pro Val Ser Gly Tyr Asp Phe Gly Val
305          290          295          300
307 gat aca agt cct ggg acc att cgc gcc ctg ccc aac gcc gca acg gct      960
308 Asp Thr Ser Pro Gly Thr Ile Arg Ala Leu Pro Asn Ala Ala Thr Ala
309 305          310          315          320
311 ggc ctc ttc gag gcg gcc gtt gag gcc act gag gaa gcg att gtt tcc      1008
312 Gly Leu Phe Glu Ala Val Glu Ala Thr Glu Glu Ala Ile Val Ser
313          325          330          335
315 gcg ctt gtc cac gcc gac acc tgc acc ggg atc gat gac agg gtt gcc      1056
316 Ala Leu Val His Ala Asp Thr Cys Thr Gly Ile Asp Asp Arg Val Ala
317          340          345          350
319 tat ggg ttg gag gcg gct cga ctt gct cgt tca att tcg gaa tat cga      1104
320 Tyr Gly Leu Glu Ala Ala Arg Leu Ala Arg Ser Ile Ser Glu Tyr Arg
321          355          360          365
323 ggc acc cag ctg tat ccg gag aaa gtg tcg gat tcc cat ctt gaa cga      1152
324 Gly Thr Gln Leu Tyr Pro Glu Lys Val Ser Asp Ser His Leu Glu Arg
325          370          375          380
327 agg agc cag ccg tga      1167
328 Arg Ser Gln Pro
329 385
332 <210> SEQ ID NO: 4
333 <211> LENGTH: 16

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/587,085

DATE: 08/03/2006  
TIME: 09:14:08

Input Set : A:\5404-161 Sequence Listing.txt  
Output Set: N:\CRF4\08032006\J587085.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; N Pos. 2,11

Seq#:5; N Pos. 6,15

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/587,085

DATE: 08/03/2006

TIME: 09:14:08

Input Set : A:\5404-161 Sequence Listing.txt

Output Set: N:\CRF4\08032006\J587085.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No

L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:351 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0

L:373 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0